NEPŘÍMÁ ORDINAČNÍ ANALÝZA

PŘEHLED METOD ORDINAČNÍ ANALÝZY

	raw-data-based (založené na primárních datech)		distance-based (založené na distanční matici)
	linear (lineární)	unimodal (unimodální)	
unconstrained (nepřímé)	PCA (analýza hlavních komponent)	CA, DCA (korespondenční a detrendovaná korespondenční analýza)	PCoA (analýza hlavních koordinát) NMDS (nemetrické mnohorozměrné škálování)
constrained (přímé)	RDA (redundanční analýza)	CCA (kanonická korespondenční analýza)	db-RDA (redundanční analýza založená na distanční matici)

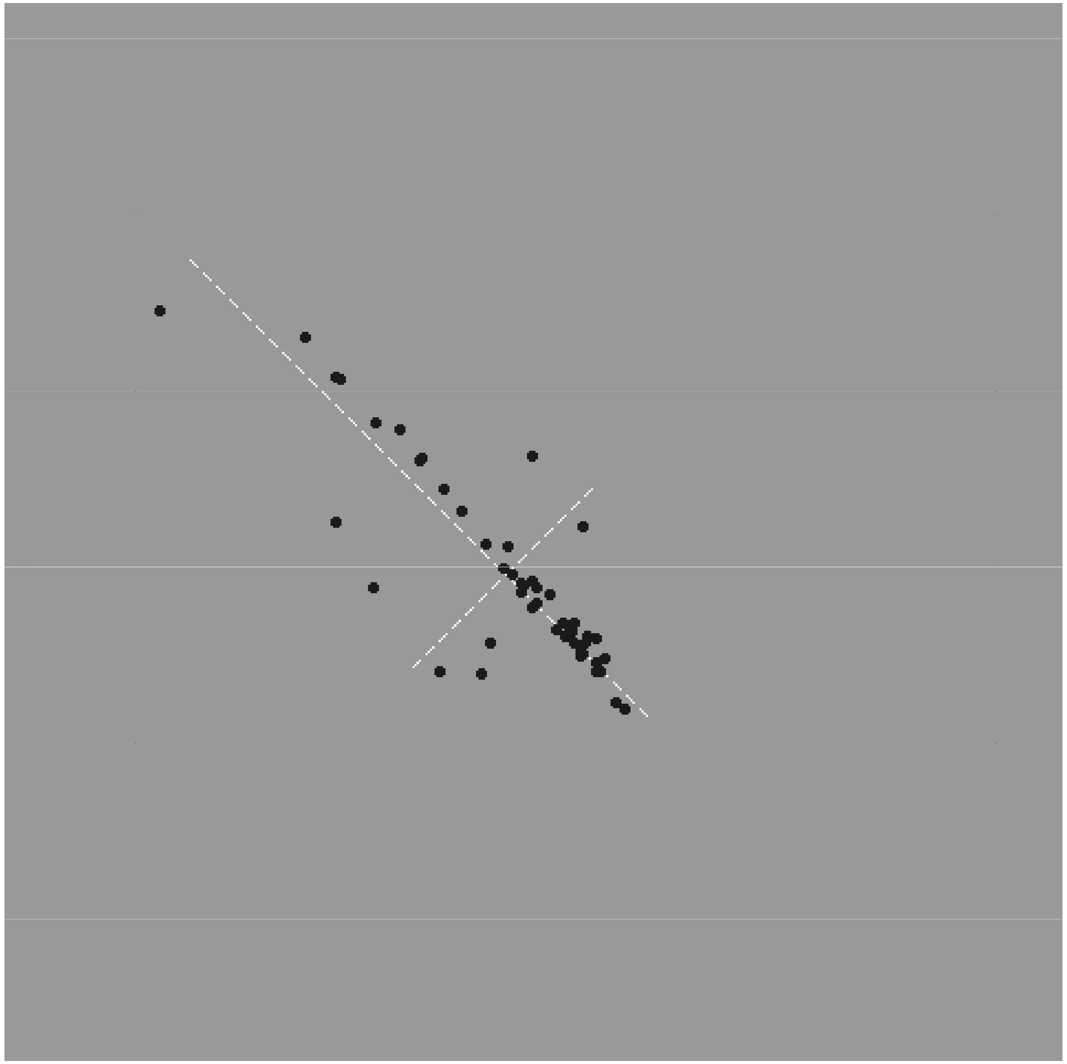
NEPŘÍMÁ EIGENVALUE-BASED ORDINACE

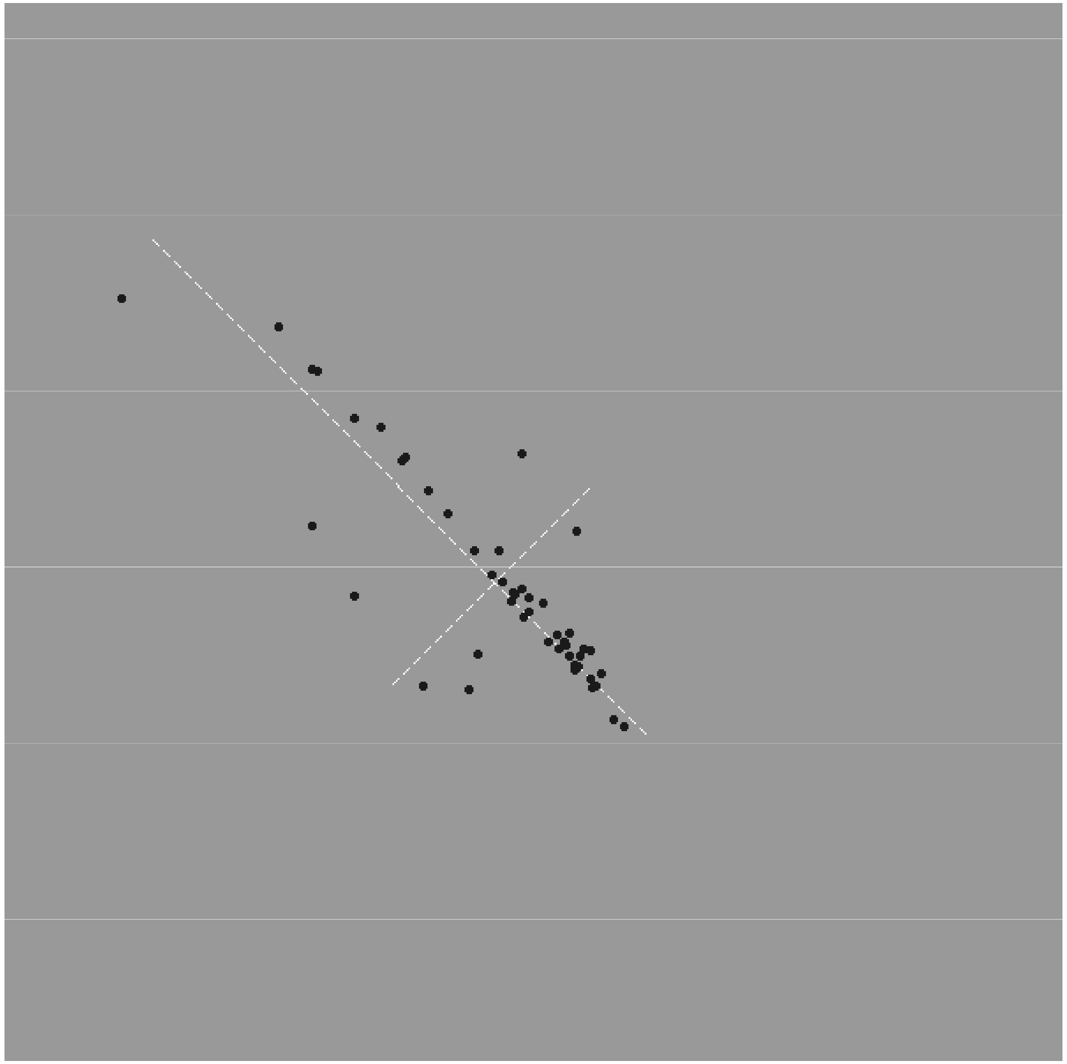
PRINCIP

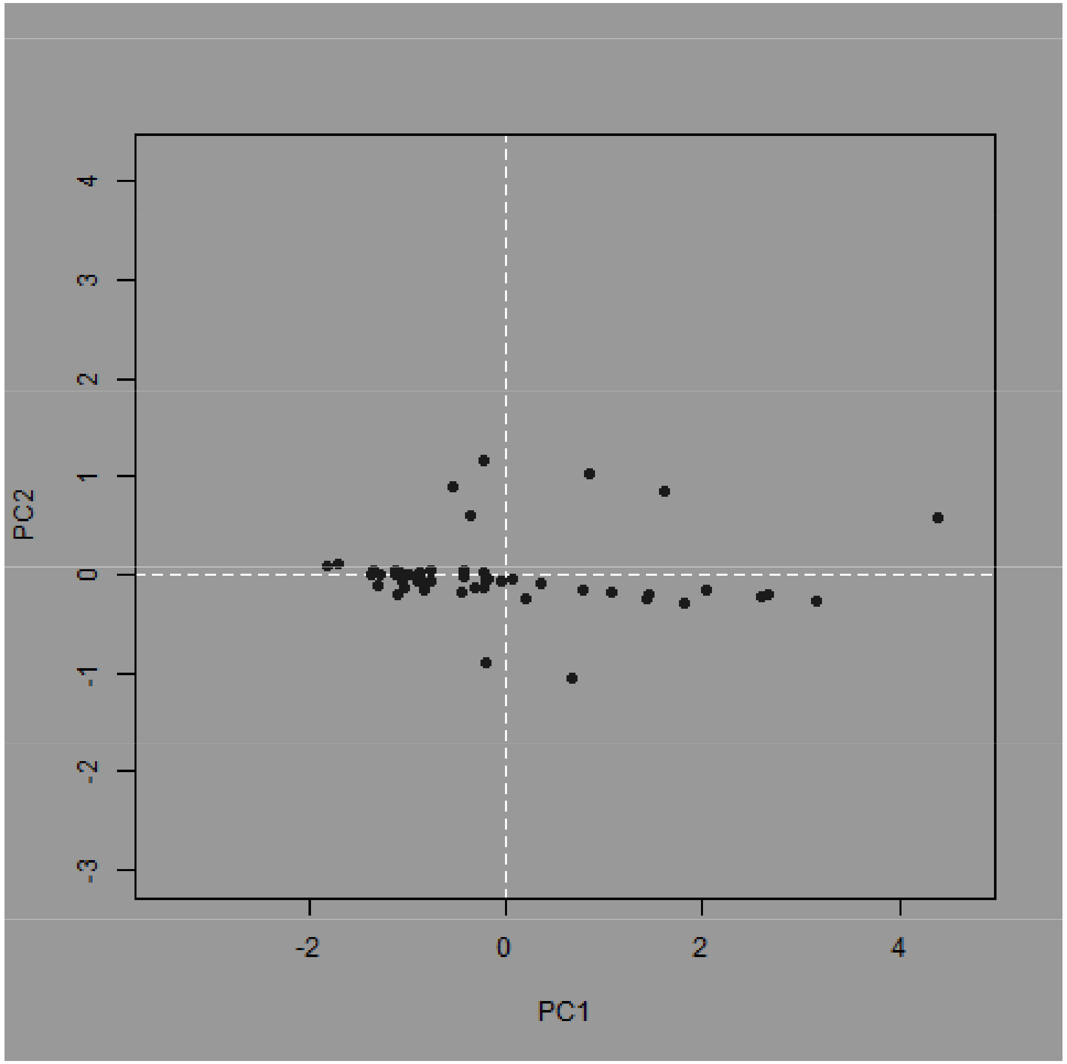
- hledání skrytých proměnných (gradientů), které nejlépe reprezentují chování všech druhů
  - seřazení vzorků podél těchto gradientů -> skóre vzorků (*sample scores*) na **ordinačních osách** (*ordination axes*)
  - odhad odpovědi (PCA, PCoA) nebo optima (DCA) jednotlivých druhů na osách (*species scores*)
- Důležitost ordinačních os (množství vysvětlené variability) klesá od první dál
- Ordinační osy jsou na sobě lineárně nezávislé – vždy přidávají novou informaci

PCA – Analýza hlavních komponent
Principal Component Analysis

PCA – Analýza hlavních komponent
Principal Component Analysis







PC2
4
3
2
1
0
-1
-2
-3
-2
0
2
4
PC1

PCA – VÝSTUP

- Eigenvalues – množství variability zachycené danou osou
- podíl vysvětlené variability – analog R^2
  - podíl eigenvalue ku sumě všech eigenvalues
- celková variabilita = suma eigenvalues = suma variancí všech proměnných
- 1. osa vysvětluje 92% variability
  - 1.8464 / (1.8464 + 0.1536)=
  - 1.8464 / 2 = 0.92

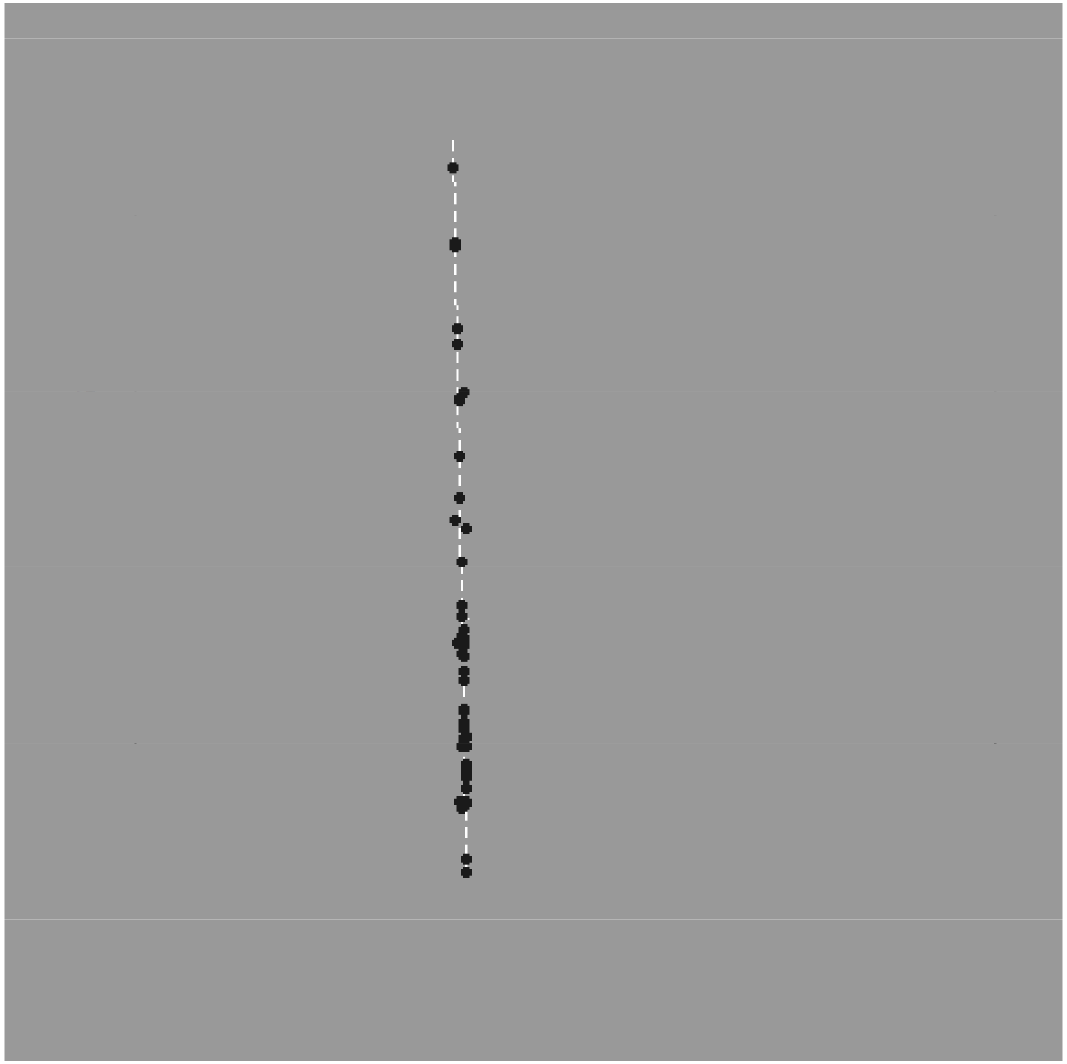
$$\frac{\lambda_k}{\sum_{k=1}^{p} \lambda_k}$$

```
                 Inertia          Rank
Total            2
Unconstrained    2                2
Inertia is variance
Eigenvalues for unconstrained axes:
PC1    PC2
1.8464 0.1536
```

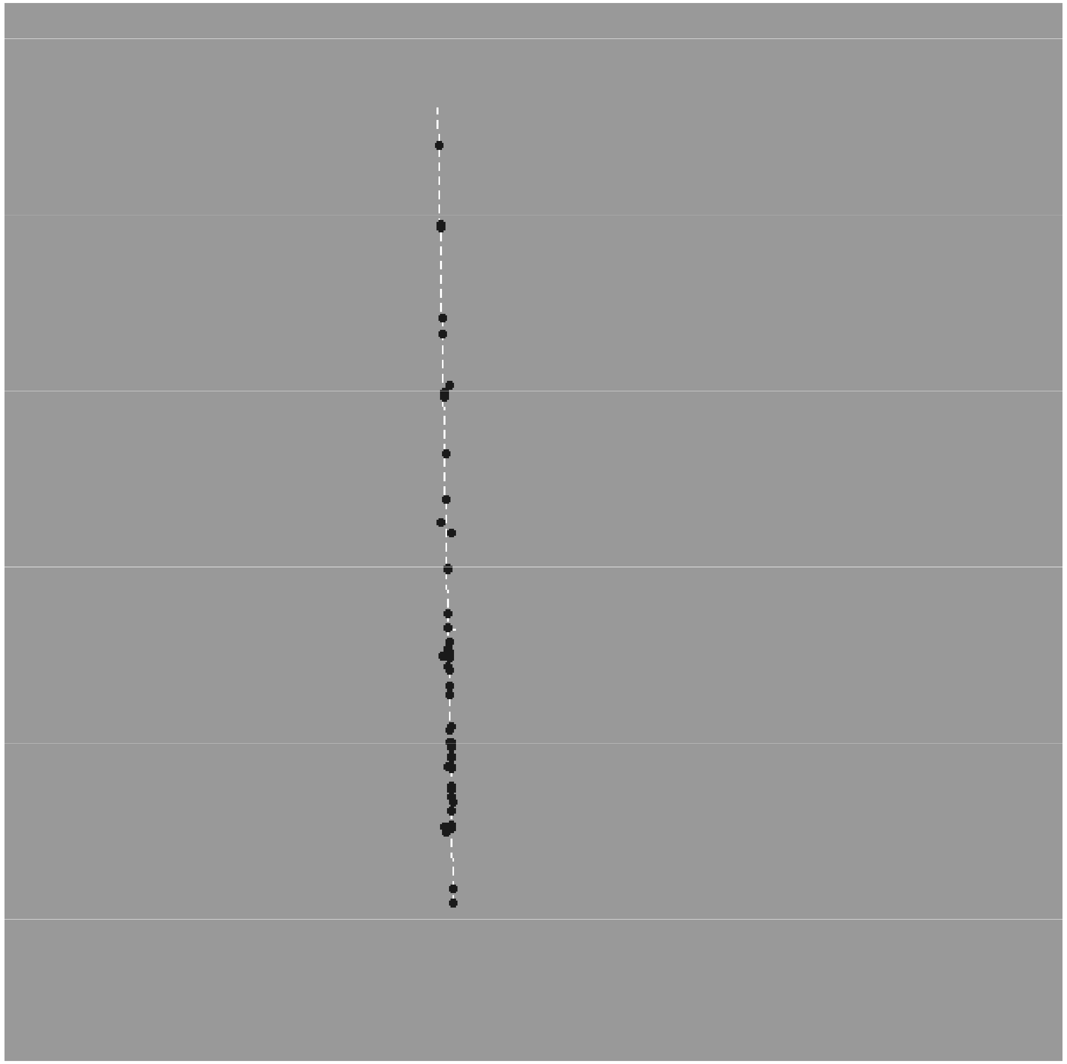
PCA – PRINCIPY

- vliv proměnné na výsledek PCA je úměrný podílu variability proměnné ku celkové variabilitě
- při standardizaci vliv každé proměnné stejný
  - variance všech proměnných = 1
- rotace PCA založena na eigenanalýze asociační matice
  - buď kovarianční matice (pokud data nejsou standardizována)
  - nebo korelační matice, pokud jsou standardizována

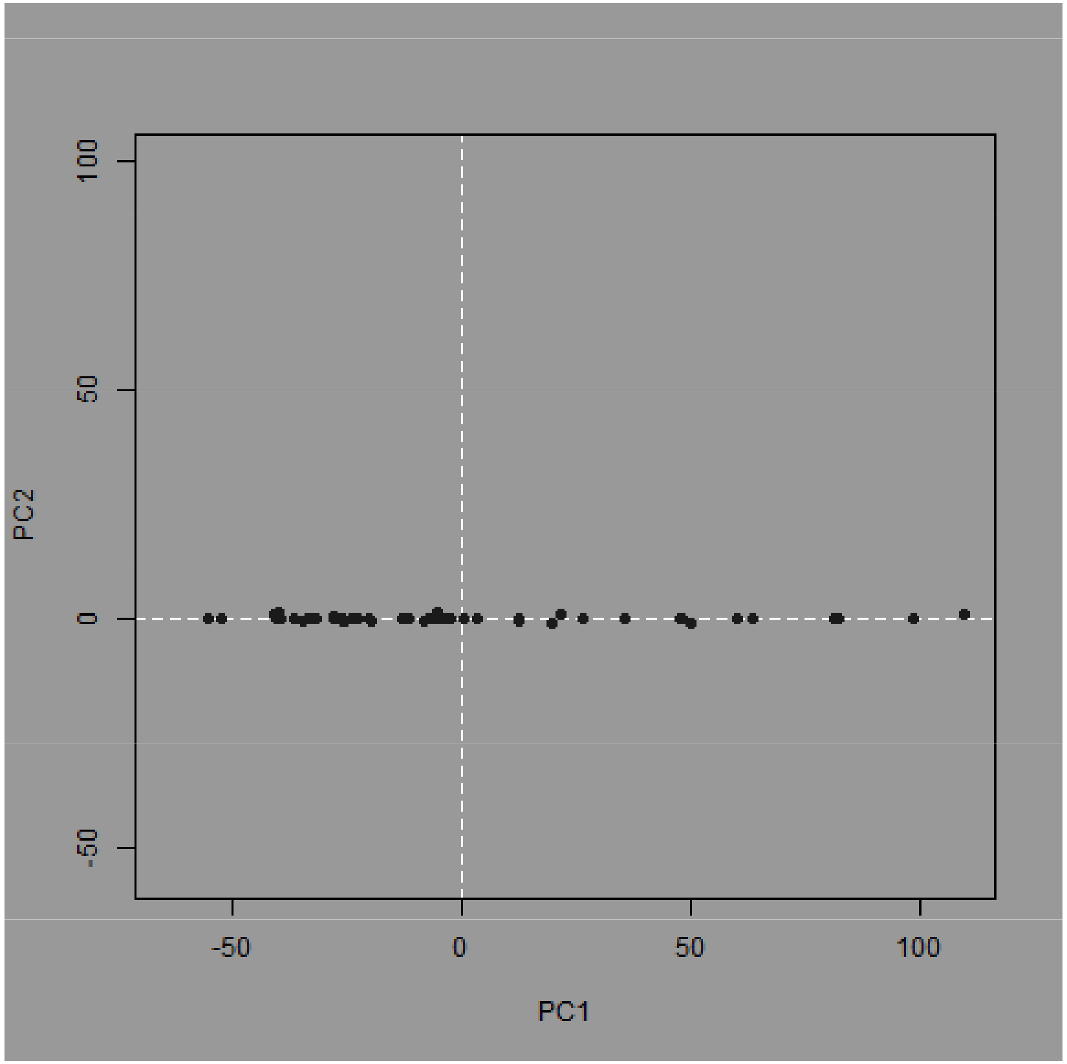
PCA NA NESTANDARDIZOVANÝCH DATECH



PCA NA NESTANDARDIZOVANÝCH DATECH

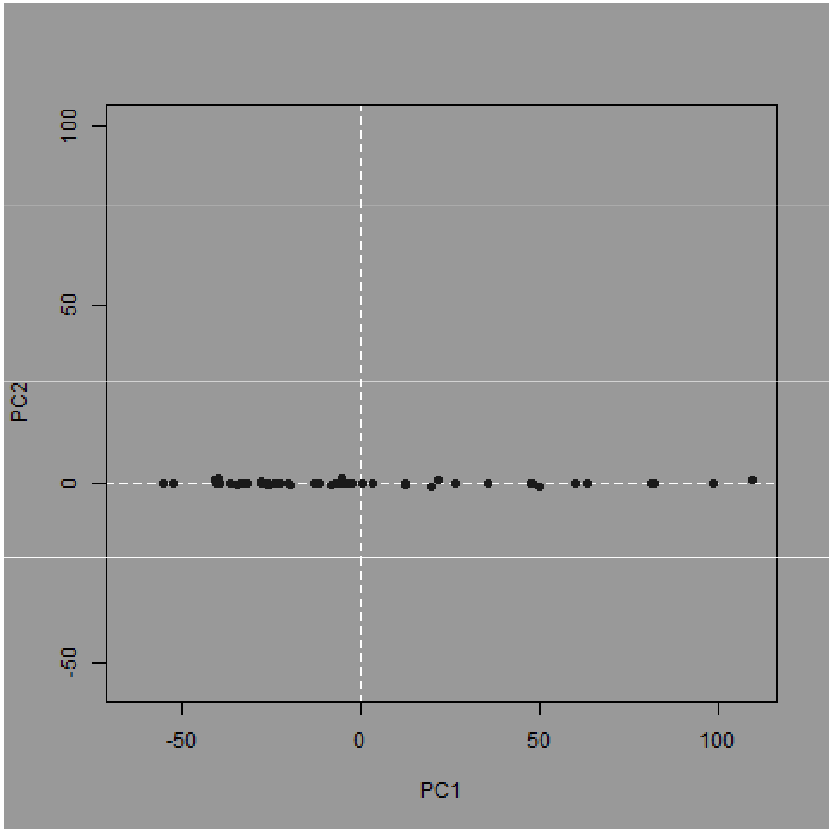


PCA NA NESTANDARDIZOVANÝCH DATECH



PCA NA NESTANDARDIZOVANÝCH DATECH

- 1. osa vysvětluje 99.9% variability

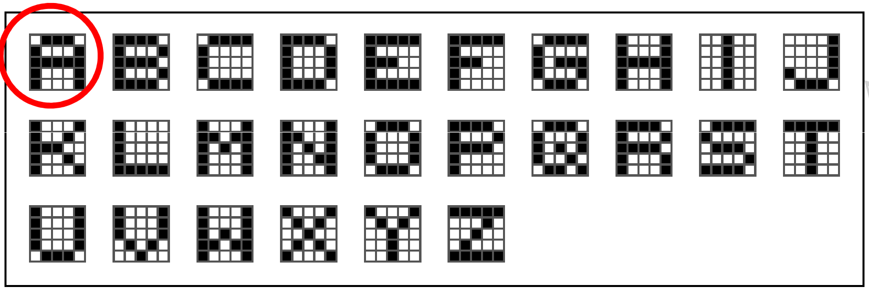


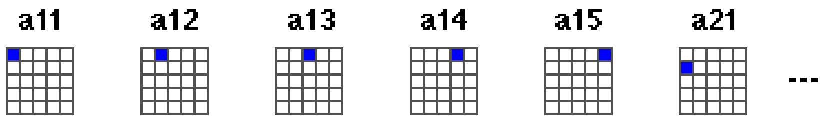
	Inertia	Rank
Total	1672	
Unconstrained	1672	2

Inertia is variance
Eigenvalues for unconstrained axes:

PC1	PC2
1671.9	0.2

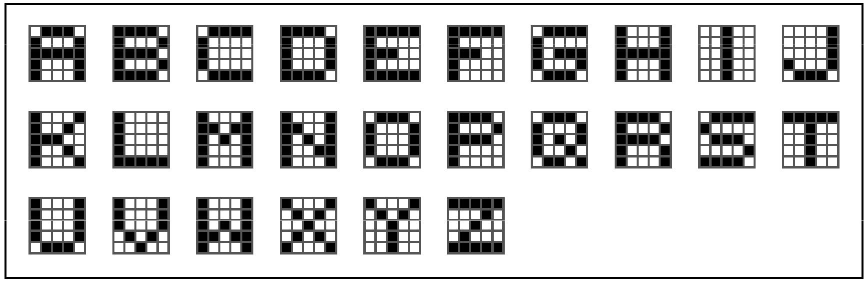
Příklad: rozeznávání písmen v analýze obrazu pomocí PCA

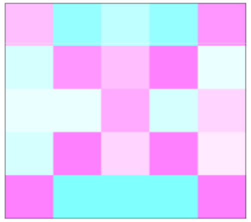




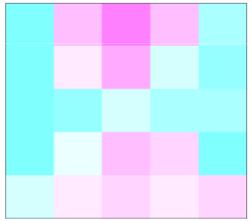
	a11	a12	a13	a14	a15	a21	a22	a23	a24	a25	a31	a32	a33	a34	a35	a41	a42	a43	a44	a45	a51	a52	a53	a54	a55
A	0	1	1	1	0	1	0	0	0	1	1	1	1	1	1	1	0	0	0	1	1	0	0	0	1
B	1	1	1	1	0	1	0	0	0	1	1	1	1	1	0	1	0	0	0	1	1	1	1	1	0
C	0	1	1	1	1	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	1	1	1	1
D	1	1	1	1	0	1	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	1	1	1	0
E	1	1	1	1	1	1	0	0	0	0	1	1	1	0	0	1	0	0	0	0	1	1	1	1	1
F	1	1	1	1	1	1	0	0	0	0	1	1	1	0	0	1	0	0	0	0	1	0	0	0	0
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
X	1	0	0	0	1	0	1	0	1	0	0	0	1	0	0	0	1	0	1	0	1	0	0	0	1
Y	1	0	0	0	1	0	1	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0
Z	1	1	1	1	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	1	1	1	1

Inspired by work of François Labelle (http://www.cs.mcgill.ca/~sqrt/dimr/dimreduction.html)

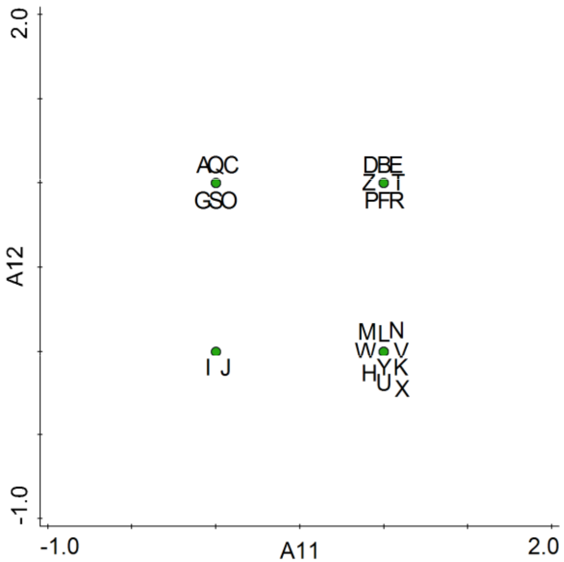




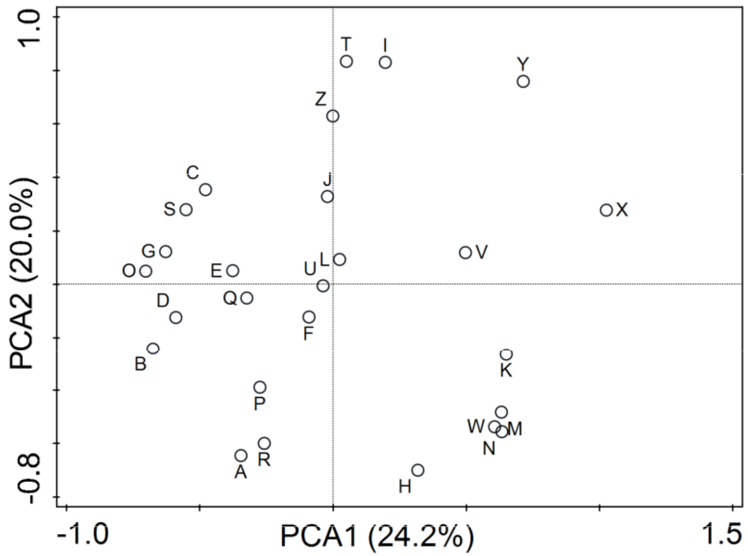
PCA1
(O-X)



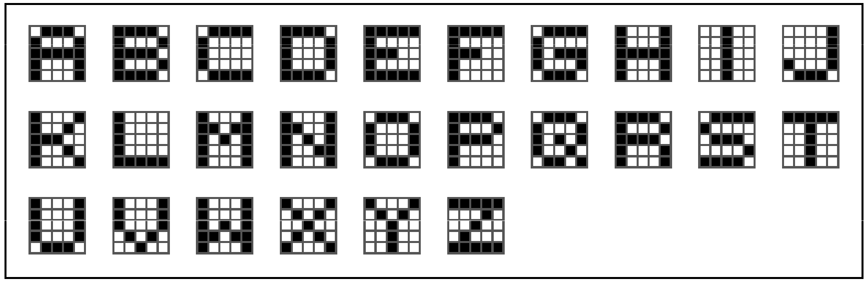
PCA2
(H-I)

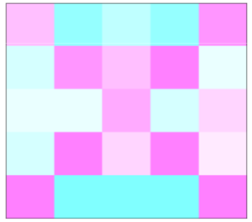


vztah proměnných A11
a A12

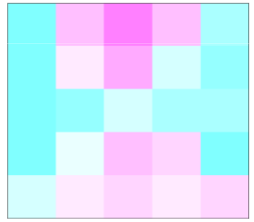


výsledek PCA
(1. a 2. PCA osa)

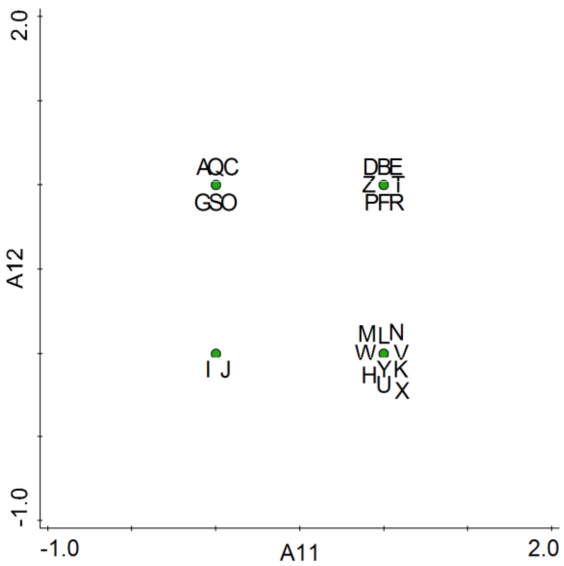




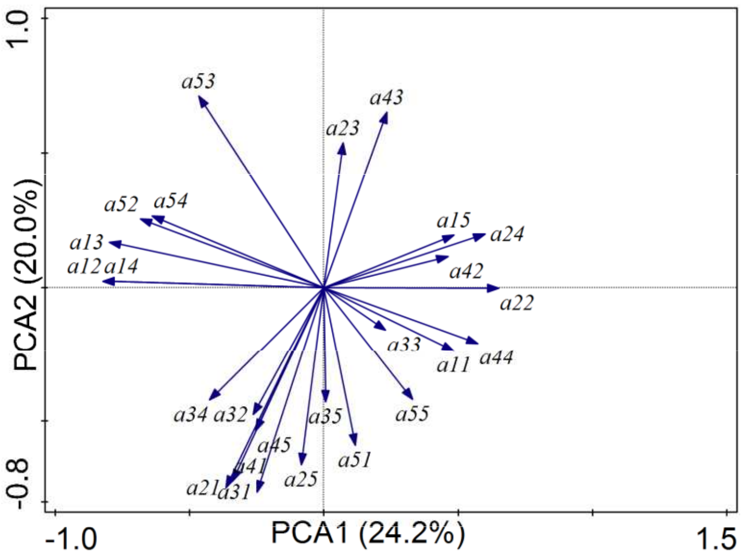
PCA1
(O-X)



PCA2
(H-I)



vztah proměnných A11
a A12

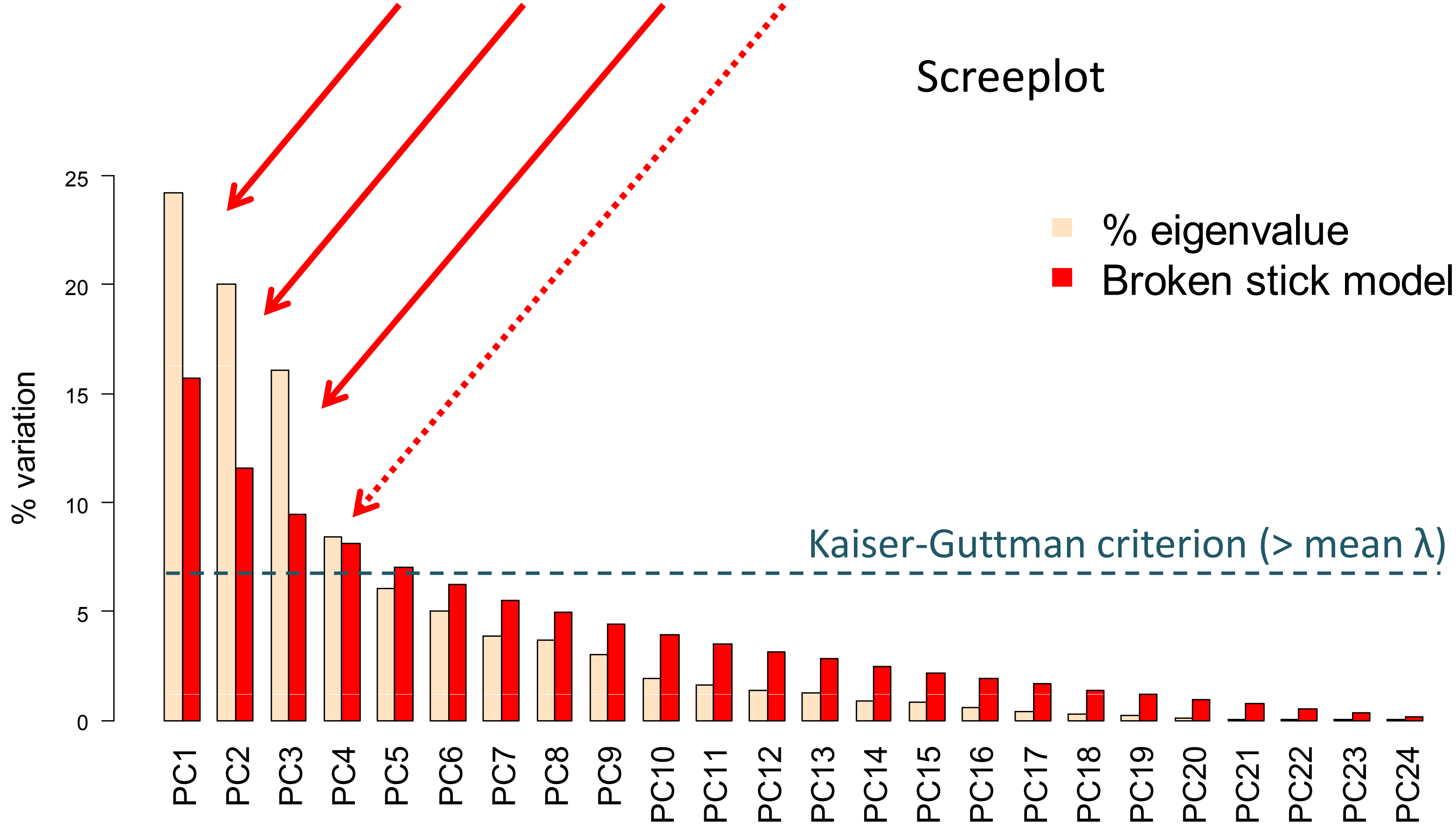


výsledek PCA
(1. a 2. PCA osa)

KTERÉ OSY PCA INTERPRETOVAT?
KTERÉ JSOU DŮLEŽITÉ?

Summary Table:

Statistic	Axis 1	Axis 2	Axis 3	Axis 4	Axis 5	Axis 6	Axis 7	Axis 8	...	Axis 23	Axis 24
Eigenvalues	0.242	0.2002	0.1608	0.0843	0.0608	0.0501	0.0389	0.0369	...	0.0002	0.0001
Explained variation (cumulative)	24.2	44.22	60.3	68.73	74.81	79.82	83.71	87.4	...	99.99	100



BROKEN-STICK MODEL (EXPECTATIONS)

model. Frontier (1976) and Legendre and Legendre (1983) provide a table of eigenvalues based on the broken-stick distribution, but the solution is easily calculated as:

$$b_k = \sum_{i=k}^{p} \frac{1}{i},$$

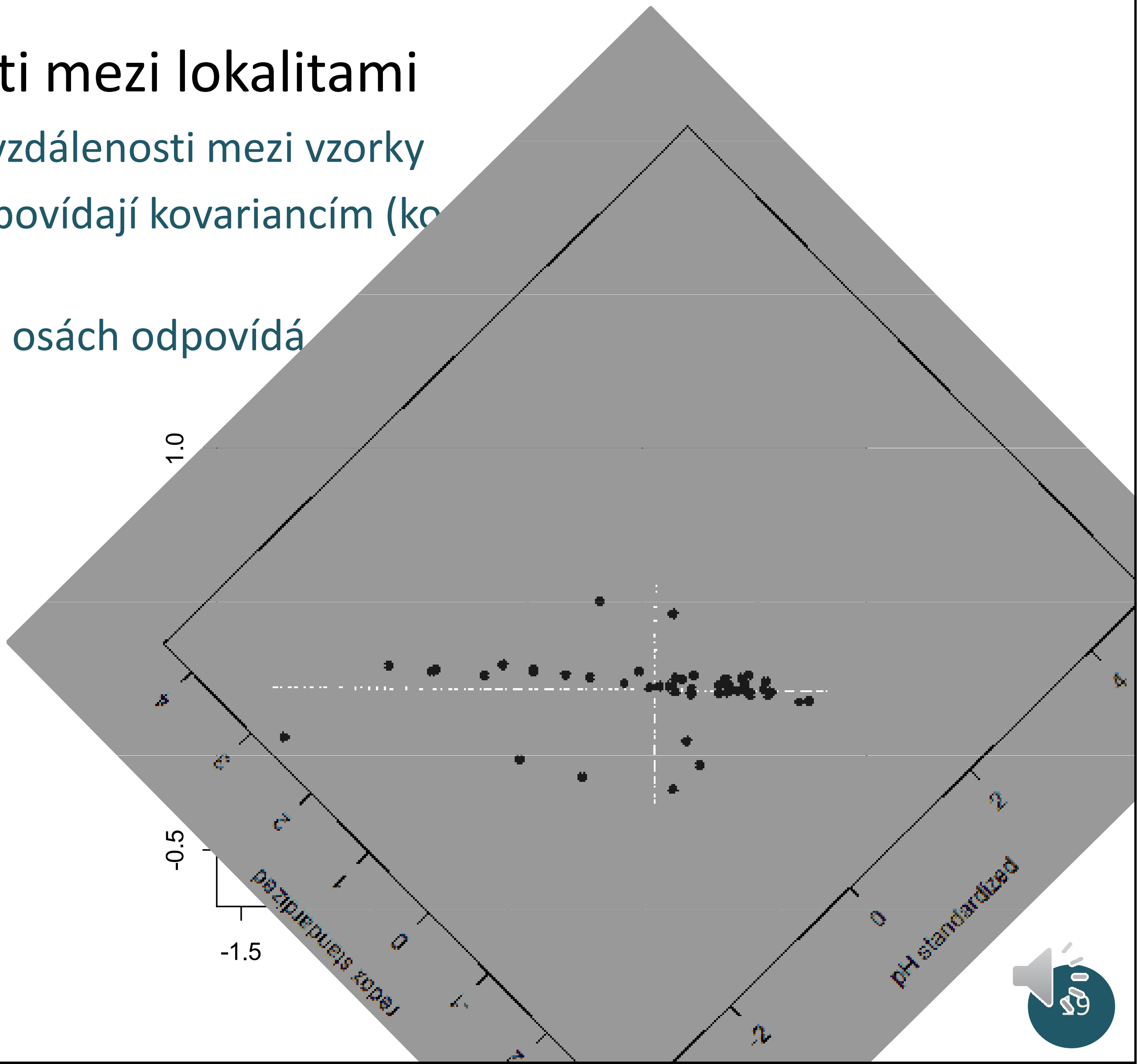
where p is number of variables and b_k is the size of the eigenvalue for the k^{th} component under the broken-stick model.

- 1. osa = 1/1+1/2+1/3+1/4+ …. 1/n
- 2. osa = 1/2+1/3+1/4+1/5 …. 1/n
- kde *n* je celkový počet os

PCA BIPLOT – ŠKÁLOVÁNÍ OS (1)

1 (sites)

- zaměření na odlišnosti mezi lokalitami
  - zachovány euklidovské vzdálenosti mezi vzorky
  - úhly mezi šipkami neodpovídají kovariancím (ko proměnných
  - variance skóre lokalit na osách odpovídá zachycené variability)

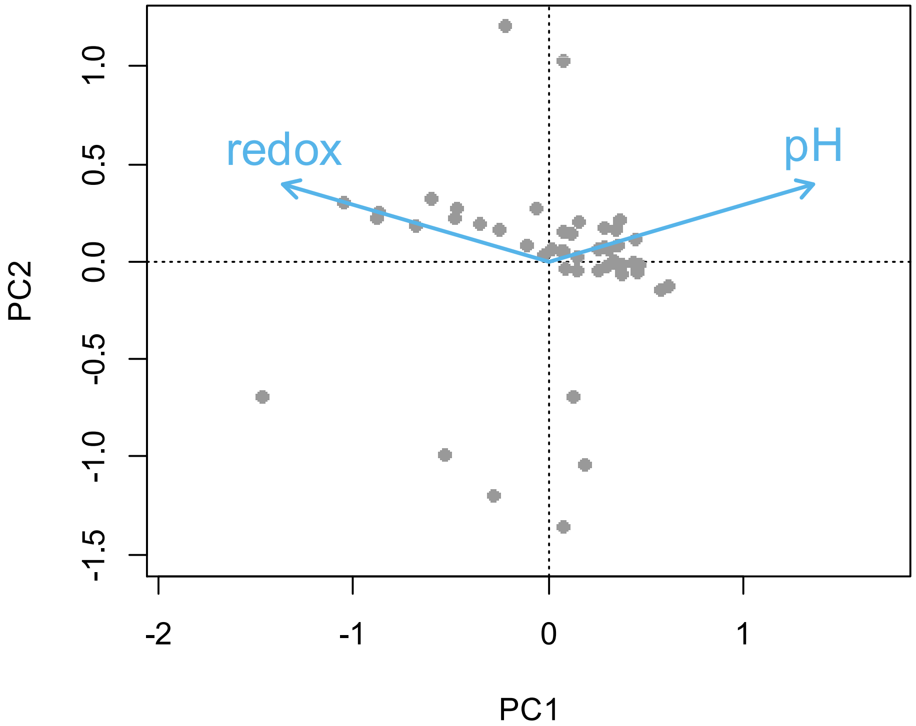


- `scale = "sites"` (nebo 1) v **R**

PCA BIPLOT – ŠKÁLOVÁNÍ OS (2)

2 (species)

- zaměření na vztahy mezi proměnnými
  - úhly mezi šipkami odpovídají kovariancím (korelacím) mezi proměnnými
  - vzdálenosti mezi vzorky neodpovídají euklidovským vzdálenostem
  - variance skóre lokalit na osách rovna 1



- `scale = „species"` (nebo 2) v **R**